



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/580,502	05/23/2006	Gerhard Grobauer	AT 030066	2330
24737 7590 05/21/2009 PHILIPS INTELLECTUAL PROPERTY & STANDARDS P.O. BOX 3001 BRIARCLIFF MANOR, NY 10510				
EXAMINER				
ALBERTALLI, BRIAN LOUIS				
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
05/21/2009		PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/580,502

Applicant(s)

GROBAUER ET AL.

Examiner

BRIAN L. ALBERTALLI

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-17 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-5, 7-12 and 15-17 is/are rejected.
- 7) ☒ Claim(s) 6, 13 and 14 is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. ____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/CDC)
- Paper No(s)/Mail Date ____.

- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date ____.
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: ____.

DETAILED ACTION

Claim Rejections - 35 USC § 101

1. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

2. Claims 1-4 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claims 1-4 do not fall within one of the four statutory categories of invention. While the claim(s) recite a series of steps or acts to be performed, a statutory "process" under 35 USC 101 must (1) be tied to another statutory category (such as a manufacture or a machine), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. The instant claim(s) neither transform underlying subject matter nor positively recite structure associated with another statutory category, and therefore do not define a statutory process.

Claim 1 is directed to a method of transcribing an audio signal into a document. Clearly then, the method does not transform the underlying subject matter. Each step of claim 1 could be performed manually by a human, thus the process is also not tied to another statutory category. Specifically, the step of transcription of signals portions into text portions could be performed by a person writing down corresponding text while listening to a spoken audio file. The step of production of relational data could be performed by a person writing down timing information (such as the playback time) along with each word. The step of recognition of the structure of the document could be

performed by a human recognizing the utterance of a section heading or the like in the audio file. Finally, the step of depicting the recognized structure in the relational data could be performed by a human marking those words that were recognized as related to the structure.

Thus, claim 1 is not sufficiently tied to another statutory category and the claimed process is nonstatutory.

Regarding claim 2, a human could read the document and recognize structure.

Regarding claim 3, a human could read the recognized text and recognize where structure words were used.

Regarding claim 4, a human could organize the written transcription in logical groups.

3. Claims 5-7 are considered to meet the "tied-to" requirement, because each claim is explicitly or inherently tied to another statutory category.

Claim 5 recites "transcription means", which are described in the specification as an element of a device (1).

Claim 6 requires reproduction of the signal portions of the audio signal. A human cannot reasonably reproduce an audio signal (i.e. provide an exact recording and playback of an audio signal). Therefore, claim 6 is inherently tied to a particular machine or apparatus.

Claim 7 recites "synthesis means", which are described in the specification as an element of device (1).

4. Claim 15 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claim 15 is directed to "a computer program product". This language is typically statutory, however, claims 16 and 17, which depend from claim 15, recite the computer program product being stored on a computer readable medium, or run by a computer. Similarly the specification states that a "computer program product" is stored on a computer readable medium and can be loaded into the memory of a computer (page 16, lines 2-9).

It would appear, therefore, that the claimed "computer program product" is nothing more than computer software code. Computer programs claimed as computer listings *per se*, i.e., the descriptions or expressions of the programs are not physical "things." They are neither computer components nor statutory processes, as they are not "acts" being performed. Such claimed computer programs do not define any structural and functional interrelationships between the computer program and other claimed elements of a computer, which permit the computer program's functionality to be realized.

Claim Rejections - 35 USC § 112

5. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

6. Claims 1 and 8 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 1 and 8 recite a document where the document is "envisaged for the reproduction of information". The use of the term "envisaged" renders the claims indefinite. Since "envisaged" can be interpreted as a conceived *possibility*, it cannot be determined what the document may be used for and the metes and bounds of the claimed "document" cannot be determined.

For the purposes of examination, a "document" has been interpreted as a visual document comprising text, figures, etc., such as that shown in Fig. 2.

Claim Rejections - 35 USC § 102

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

8. Claims 1-4 and 8-11 are rejected under 35 U.S.C. 102(b) as being anticipated by Goldhor et al. (U.S. Patent 5,970,448).

In regard to claim 1, Goldhor et al. disclose a method for transcribing an audio signal (AS) containing signal portions (SP) into text containing text portions (TP) for a document (DO), this document (DO) being envisaged for the reproduction of information, this information corresponding at least in part to the text portions (TP)

obtained through the transcription (generate text from voice input, column 2, lines 13-16; to create textual documents, column 12, lines 41-50), this method having the steps listed below, namely:

transcription of the signal portions (SP) into text portions (TP) (Fig. 4a, generation of text from input speech, column 8, lines 24-39) and

production of relational data (RD) which represent at least one temporal relation between respectively at least one signal portion (SP) and respectively at least one text portion (TP) obtained through transcription (for each input speech event, dictation event records and text event records are generated, see Figs. 2 and 3 and column 2, lines 40-54; the records including a best match recognition element representing the textual interpretation of the input voice, column 4, lines 33-42; the waveform of the speech for the speech event, column 4, lines 47-54; and chronological relationship information to relate the timing of the speech events, column 6, lines 6-15), and

recognition of a structure of the document (DO) (in a form filling mode, a particular form field in a document is selected by voice, column 10, line 58 to column 11, line 19) and

depiction of the recognized structure of the document (DO) in the relational data (RD) (as mentioned above, all speech events generate dictation event records and text event records are generated, column 2, lines 50-54; thus, the dictation and text events used to fill in forms, column 10, lines 58-65, are included in the dictation event records and text event records).

In regard to claim 2, Goldhor et al. disclose the recognition of the structure of the document (DO) takes place through analysis of the document (DO) (the system creates a text event record for each field in a particular form document, thus allowing a user to select a field by voice, column 10, line 58 to column 11, line 8).

In regard to claim 3, Goldhor et al. disclose the recognition of the structure of the document (DO) takes place through analysis of the recognized text portions (TP) (during dictation, the system determines there is an active dictation event with a corresponding active text event (which are generated by the recognition of input voice), and the form associated with the text event is recognized, column 11, lines 9-19).

In regard to claim 4, Goldhor et al. disclose the depiction of the recognized structure of the document (DO) takes place through a logical grouping of the relational data (RD) (text events associated with each field in the form are logically associated, column 10, line 58 to column 11, line 8).

In regard to claim 8, Goldhor et al. disclose a device (Fig. 1) for transcribing an audio signal (AS) containing signal portions (SP) into text containing text portions (TP) for a document (DO), this document (DO) being envisaged for the reproduction of information, this information corresponding at least in part to the text portions (TP) obtained through the transcription (generate text from voice input, column 2, lines 13-16; to create textual documents, column 12, lines 41-50)

with transcription means for the transcription of the signal portions (SP) into text portions (TP) (Fig. 4a, generation of text from input speech, column 8, lines 24-39) and

with relational data production means which are designed for the production of relational data (RD) which represent at least one temporal relation between respectively at least one signal portion (SP) and respectively at least one text portion (TP) obtained through transcription (for each input speech event, dictation event records and text event records are generated, see Figs. 2 and 3 and column 2, lines 40-54; the records including a best match recognition element representing the textual interpretation of the input voice, column 4, lines 33-42; the waveform of the speech for the speech event, column 4, lines 47-54; and chronological relationship information to relate the timing of the speech events, column 6, lines 6-15), and

with structure recognition means which are designed for recognition of a structure of the document (DO) (in a form filling mode, a particular form field in a document is selected by voice, column 10, line 58 to column 11, line 19) and

with structure depiction means which are designed for depicting the recognized structure of the document (DO) in the relational data (RD) (as mentioned above, all speech events generate dictation event records and text event records are generated, column 2, lines 40-54; thus, the dictation and text events used to fill in forms, column 10, lines 58-65, are included in the dictation event records and text event records).

In regard to claim 9, Goldhor et al. disclose the structure recognition means (6) are realized with the aid of a first analysis stage (7) which is designed for analyzing the

document (DO) in respect of its structure (the system creates a text event record for each field in a particular form document, thus allowing a user to select a field by voice, column 10, line 58 to column 11, line 8).

In regard to claim 10, Goldhor et al. disclose the structure recognition means (6) are realized with the aid of a second analysis stage (8), which is designed for analyzing the text portions (TP) obtained in respect of a structure of the document (DO) (during dictation, the system determines there is an active dictation event with a corresponding active text event (which are generated by the recognition of input voice), and the form associated with the text event is recognized, column 11, lines 9-19).

In regard to claim 11, Goldhor et al. disclose the structure depiction means (9) are designed for the logical grouping of the relational data (RD) (text events associated with each field in the form are logically associated, column 10, line 58 to column 11, line 8).

Claim Rejections - 35 USC § 103

9. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

10. Claims 6 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Goldhor et al., in view of Holt et al. (U.S. Patent 5,960,447).

In regard to claim 6, Goldhor et al. do not disclose transcription means (2), provided for the transcription of text portions (TP), are configured depending on the recognized structure.

Holt et al. disclose a method for transcribing an audio signal wherein transcription means (2), provided for the transcription of text portions (TP), are configured depending on the recognized structure (see Fig. 5, when the focus of the recognition is in a particular field of a document, recognition is constrained to a vocabulary specific to that field, column 9, line 66 to column 10, line 12).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor et al. to configure the transcription means depending on the recognized structure of the document, because, as is recognized by those of ordinary skill in the art, constraining the vocabulary of a transcriber to a limited vocabulary of valid words increases the correct recognition rate. Thus, by modifying Goldhor et al. to configure the transcriber based on the structure of the document (i.e. the valid vocabulary for a given field), the probability of correct recognition for a particular field would increase.

In regard to claim 12, Goldhor et al. do not disclose the transcription means (2), can be configured depending on the recognized structure.

Holt et al. disclose a system for transcribing an audio signal wherein transcription means (2), can be configured depending on the recognized structure (see Fig. 5, when the focus of the recognition is in a particular field of a document, recognition is constrained to a vocabulary specific to that field, column 9, line 66 to column 10, line 12).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor et al. to configure the transcription means depending on the recognized structure of the document, because, as is recognized by those of ordinary skill in the art, constraining the vocabulary of a transcriber to a limited vocabulary of valid words increases the correct recognition rate. Thus, by modifying Goldhor et al. to configure the transcriber based on the structure of the document (i.e. the valid vocabulary for a given field), the probability of correct recognition for a particular field would increase.

11. Claims 7 and 15-17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Goldhor et al., in view of Reynar et al. (U.S. Patent 6,446,041).

In regard to claim 7, Goldhor et al. do not disclose further text portions (TP'), produced in addition to the text portions (TP) obtained through the transcription of the audio signal (AS), which further text portions (TP') exist adjacent to the text portions (TP) obtained through the transcription of the audio signal (AS) in the document (DO), are reproduced with the aid of speech that can be created by synthesis means, and

wherein if necessary the reproduction of the audio signal (AS) is interrupted during the reproduction of the further text portions (TP').

Reynar et al. disclose a method of speech transcription, wherein further text portions (TP'), produced in addition to the text portions (TP) obtained through the transcription of the audio signal (AS) (a user edits a dictation document by manually adding new text portions, column 4, lines 62-67), which further text portions (TP') exist adjacent to the text portions (TP) obtained through the transcription of the audio signal (AS) in the document (DO) (see Fig. 3B, a manually edited portion 320 is added to the document adjacent to dictated portions, column 11, lines 1-17), are reproduced with the aid of speech that can be created by synthesis means, and wherein if necessary the reproduction of the audio signal (AS) is interrupted during the reproduction of the further text portions (TP') (during playback of a selected portion of a document, if speech input is associated with a given word, the speech input is played back, column 11, lines 54-61; if there is no speech input associated with a given word, the text is synthesized by a TTS, column 12, lines 9-18 and lines 28-31).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor et al. to reproduce further obtained text portions with the aid of speech that could be created by speech synthesis means, because this allows the entire document to be played back, even if there is no dictation audio associated with a given text portion, which minimizes user confusion caused by skipping non-dictated text portions, as taught by Reynar et al. (column 3, lines 58-67).

In regard to claim 15, Goldhor et al. do not explicitly disclose a computer program product which is suitable for the transcription of an audio signal (AS) and which can be loaded directly into a memory of a computer and includes software code sections, wherein with the computer, the method as claimed in claim 1 can be executed when the computer program product is run on the computer.

Reynar et al. disclose a computer program product which is suitable for the transcription of an audio signal (AS) and which can be loaded directly into a memory of a computer and includes software code sections (program modules loadable into a memory 122 of computer 100, see Fig. 1, column 6, lines 32-44 and column 7, lines 52-63).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor et al. to implement the method of claim 1 as a computer program product loadable into the memory of a computer, because, as is notoriously well-known by those of ordinary skill in the art, computer program products (program modules) define the logic that allow a computer to perform the computer program product's functionality and allow the computer to realize the logic of the computer program product.

In regard to claim 16, Goldhor et al. do not explicitly disclose the computer program product is stored on a computer-readable medium.

Reynar et al. disclose the computer program product is stored on a computer-readable medium (computer readable media, column 6, lines 53-62).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor to store the computer program product on a computer-readable medium, because, as is notoriously well-known by those of ordinary skill in the art, such computer readable mediums define structural and functional interrelationships between the computer program and other claimed elements of a computer, which permit the computer program's functionality to be realized.

In regard to claim 17, Goldhor et al. do not explicitly disclose a computer with a computing unit and an internal memory, which runs the computer program product as claimed in claim 15.

Reynar et al. disclose a computer with a computing unit and an internal memory, which runs the computer program product (computer 100 including memory 122, column 6, lines 32-44).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Goldhor et al. to provide a computer with a memory to run the computer program product of 15, because, as is notoriously well-known by those of ordinary skill in the art, providing a computer would allow the computer program product to be executed, and thus the functionality of the computer program product would be realized.

Allowable Subject Matter

12. Claims 6, 13, and 14 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is a statement of reasons for the indication of allowable subject matter: Regarding claims 6 and 13, Goldhor et al. do not disclose or suggest reproducing signal portions at the same time as a visual emphasis of the transcribed text portions. Although the prior art teaches reproducing signal portions at the same time as a visual emphasis of the transcribed text portions (see, e.g. Wutte U.S. Patent 6,792,409), there is no teaching or suggestion to take into account the recognized structure of a document while reproducing signal portions at the same time as a visual emphasis of the transcribed text portions.

Conclusion

13. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Lucas et al. (U.S. Patent 6,834,264) disclose a method for transcriptionists to fill out structured documents from dictated input. Hollerbauer discloses storing dictated audio data with links to corresponding text portions. Forbes (U.S. Patent 7,444,285) discloses a method of inserting speech recognition results into document templates. Brais et al. (U.S. Patent 5,995,936) disclose a system that links sound and image data to recognized text portions. Mitchell et al. (U.S. Patent

5,857,099) disclose a method of linking audio data and text data. Groner et al. (U.S. Patent 6,813,603) disclose a system for populating form fields using dictation output.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to BRIAN L. ALBERTALLI whose telephone number is (571)272-7616. The examiner can normally be reached on Monday-Thursday, 8 AM to 6:30 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.